AAS 94-169

A CONTINUOUS SQUARE ROOT INFORMATION FILTER-SMOOTHER WITH DISCRETE DATA UPDATE

James K. Miller[†]

Data filters may be separated into two categories being continuous filters or discrete filters. Continuous data filters are described by a matrix differential equation or Ricatti equation and discrete filters are obtained by solving the continuous equations over some finite time interval.

A continuous matrix differential equation for the square root information (SRIF) filter is derived and a computer algorithm developed to implement this filter. The filter algorithm performs the mapping of state and process noise by numerical integration of the SRIF matrix and admits data via a discrete least square update. A smoothing algorithm is also described. For comparison, the matrix differential equations for the covariance filter, the information filter, and the square root covariance filter are also derived.

Accuracy, computational efficiency, memory requirements and simplicity of design are compared with other filter algorithms. A preliminary evaluation of these competing criteria reveals no significant differences. Computational efficiency seems to favor discrete filters since these have been developed to a high degree of efficiency. The continuous SRIF filter is expected to out perform discrete filters with regard to accuracy and memory requirements because there is no need to compute the state transition matrix and error control may be placed directly on the elements of the SRIF matrix. Simplicity of design also favors the continuous SRIF filter particularly when the same numerical integration algorithm used to propagate the state is also used to integrate the SRIF matrix.

INTRODUCTION

A data filter processes data in order to obtain an estimate of parameters that are related to the data by a mathematical model. Data filters exist in many forms and use the covariance of the state parameters, or some equivalent representation, along with the measurements and a simulation of the measurements including partial derivatives, to obtain the desired estimate. The covariance matrix of the state may be represented by its inverse or information matrix or square root factorizations of either of these matrices.

Data filters may be separated into two categories being continuous filters or discrete filters. Continuous data filters are described by a matrix differential equation or Ricatti equation and discrete filters are obtained by solving the continuous equations over some finite time interval. Discrete data filters have been developed to a high degree of efficiency and are used almost exclusively for processing data.

Continuous data filters have an advantage of inherent simplicity. The complete elimination of the need to compute a state transition matrix results in a simple algorithm involving numerical integration. The continuous filter thus promises increased accuracy since the error control may be placed directly on the covariance, or representation of the covariance, rather than on the state transition matrix.

In this paper, a continuous data filter is developed that is based on the square root factorization of the information matrix. A smoothing algorithm is also described.

[†] Member of the Technical Staff, Navigation Systems Section, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, Associate Fellow AIAA

SYSTEM DYNAMICS

The system dynamics may be described as a linear perturbation of a reference function of the state variables. For example, given the nominal trajectory of a spacecraft described by the function $x(t)$ and a perturbation of the state (δx) at the initial epoch (t_0), the perturbed trajectory is described by

$$x(t) = x(t) + \Phi(t, t_0)\, \delta x(t_0) \tag{1}$$

where the state transition matrix (Φ) is given by

$$\Phi = \frac{\partial x(t)}{\partial x(t_0)}$$

The state transition matrix may be obtained as a solution of the following differential equation or by numerical integration.

$$\frac{\partial \dot{x}(t)}{\partial x(t_0)} = \frac{\partial \dot{x}(t)}{\partial x(t)} \frac{\partial x(t)}{\partial x(t_0)}$$

$$\dot{\Phi}(t, t_0) = F\, \Phi(t, t_0) \tag{2}$$

where

$$F = \frac{\partial \dot{x}(t)}{\partial x(t)}$$

The above differential equation describing the evolution of the state variation may be generalized to include other parameters and process noise.

$$\dot{X} = F X + G \Omega \tag{3}$$

where G is the mapping of Ω, the process noise. Here, the δ's have been dropped and the variation δx is represented by X. The state vector variation X may be generalized to include constant parameters (y) and stochastic parameters (p) as well as the dynamic state variables (x). The process noise (Ω) contains white noise (ω) on the stochastic parameters. Thus we have

$$X = \begin{Vmatrix} p \\ x \\ y \end{Vmatrix} \qquad \Omega = \begin{Vmatrix} \omega \\ 0 \\ 0 \end{Vmatrix} \tag{4}$$

The stochastic parameters (p) provide a means of introducing process noise into the state variables. These are defined by scalar differential equations of the form

$$\dot{p}_i = -\frac{1}{\tau_i} p_i + \omega_i \tag{5}$$

where τ_i is the correlation time and ω_i is the white noise associated with the i'th stochastic parameter. Thus, white noise is introduced directly to the parameter p and indirectly to the state via the mapping matrix F.

An estimate of the state is obtained from a mathematical model of the system dynamics that include measurements processed by a data filter. The "best" estimate of the variation of the state (X) is described by the following equations,

$$\dot{X} = F X + G \Omega + K Z \tag{6}$$

$$Z = Z - H X \tag{7}$$

$$H = \frac{\partial Z}{\partial X} \tag{8}$$

where K is the Kalman gain, $\hat{\Omega}$ represents an estimate of the process noise, Z are the actual measurements and H is the matrix of data partials. The Kalman gain is computed as a function of the measurement error, the data partials and the state error covariance (P). Thus, in order to obtain a complete set of equations that would enable the computation of the estimated state we need an equation for the Kalman gain and an equation for evolving P as a function of time.

DERIVATION OF CONTINUOUS FILTER EQUATIONS

The covariance of the state estimate is defined by the expected value represented by

$$P = E\left\{X\, X^T\right\} \tag{9}$$

As an alternative, we may compute the information matrix (Λ), the square root of the covariance (S), or the square root of the information matrix (R). The equations that define these matrices are given by

$$P = \Lambda^{-1} \tag{10}$$

$$P = SS^T \tag{11}$$

$$P^{-1} = R^T R \tag{12}$$

Thus, we are interested in obtaining differential equations of the form

$$\dot{P} = \dot{P}_m + \dot{P}_q + \dot{P}_d \tag{13}$$

$$\dot{\Lambda} = \dot{\Lambda}_m + \dot{\Lambda}_q + \dot{\Lambda}_d \tag{14}$$

$$\dot{S} = \dot{S}_m + \dot{S}_q + \dot{S}_d \tag{15}$$

$$\dot{R} = \dot{R}_m + \dot{R}_q + \dot{R}_d \tag{16}$$

where the subscript m refers to the mapping terms, the subscript q refers to process noise terms, and the subscript d refers to the data update terms.

Mapping Term

The evolution of the covariance as a function of time[1] may be obtained by mapping the state covariance obtained at some epoch (t_0) to some time in the future (t) with the state transition matrix.

$$P(t) = \Phi(t,t_0)\, P(t_0)\, \Phi(t,t_0)^T \tag{17}$$

Taking the derivative with respect to time we obtain

$$\dot{P}(t) = \dot{\Phi}(t,t_0)\, P(t_0)\, \Phi(t,t_0)^T + \Phi(t,t_0)\, P(t_0)\, \dot{\Phi}(t,t_0)^T \tag{18}$$

Since the state transition matrix is obtained by integrating

$$\dot{\Phi}(t,t_0) = F(t)\, \Phi(t,t_0) \tag{19}$$

we obtain after substitution

$$\dot{P}_m = FP + PF^T \tag{20}$$

Process Noise Term

In the covariance matrix differential equation, process noise enters as a simple addition to the covariance. Thus we have

$$P(t+\Delta t) = P(t) + G \Delta Q G^{T} \tag{21}$$

where ΔQ is the covariance of the process noise admitted over the time interval Δt and

$$\Delta Q = Q \Delta t$$

where Q is the rate of accumulation of process noise. Thus, in the continuum we have

$$\dot{P}_q = \lim_{\Delta t \to 0} \left\{ \frac{P(t+\Delta t) - P(t)}{\Delta t} \right\} = G\, Q\, G^{T} \tag{22}$$

Data Update Term

The discrete covariance update may be obtained assuming an additional measurement H_{n+1} is added to a previously determined estimate based on measurements H_n with covariance P_n. The derivation is given in many references[2,3] that are available.

$$P_{n+1} = \left(H_n^{T} \Delta W_n H_n + H_{n+1}^{T} \Delta W_{n+1} H_{n+1} \right)^{-1} \tag{23}$$

In the notation used here, H_n is a matrix with n rows corresponding to the measurements and m columns corresponding to the state parameters. H_{n+1} is a row matrix of dimension m. We also have for the covariance update,

$$P_{n+1}^{-1} = P_n^{-1} + H_{n+1}^{T} \Delta W_{n+1} H_{n+1} \tag{24}$$

and since

$$\Lambda = P^{-1}$$

$$\Lambda_{n+1} = \Lambda_n + H_{n+1}^{T} \Delta W_{n+1} H_{n+1} \tag{25}$$

Over the time interval Δt between measurements, information accumulates at a rate W and

$$\Delta W_{n+1} = W \Delta t \tag{26}$$

$$\Lambda_{n+1} - \Lambda_n = H_{n+1}^{T} W \Delta t\, H_{n+1} \tag{27}$$

Dividing by Δt and taking the limit as Δt approaches zero,

$$\dot{\Lambda}_m = H^{T} W H \tag{28}$$

we obtain a differential equation for the evolution of the information matrix due to addition of data.

Least Squares Data Update

In order to complete the filter equations, we need an algorithm for processing the measurements to obtain a best estimate of the state. The discrete form of the Kalman update algorithm is given by[2]

$$\Delta K = \dot{P} H^{T} (\Delta V + HPH^{T}) \tag{29}$$

An equivalent expression is obtained by use of the matrix inversion lemma

$$\Delta K = \left[\hat{P}^{-1} + H^{\mathrm{T}} \Delta W\, H\right]^{-1} H^{\mathrm{T}} \Delta W \tag{30}$$

If we admit the data at a rate W over a time interval Δt we have

$$\Delta K = \left[\hat{P}^{-1} + H^{\mathrm{T}} W \Delta t\, H\right]^{-1} H^{\mathrm{T}} W \Delta t \tag{31}$$

where P is the covariance at the beginning of the interval prior to processing the data. Dividing through by Δt and taking the limit as Δt approaches zero we obtain

$$K = \lim_{\Delta t \to 0} \left\{ \frac{\Delta K}{\Delta t} \right\} = P H^{\mathrm{T}} W \tag{32}$$

a differential equation for the Kalman update.

Filter Differential Equations

Collecting the terms derived above, we have the following matrix differential equation or Ricatti equation for the covariance filter,

$$\dot{P} = FP + PF^{\mathrm{T}} + GQG^{\mathrm{T}} + \dot{P}_d \tag{33}$$

$$K = P H^{\mathrm{T}} W \tag{34}$$

and for the information filter,

$$\dot{\Lambda} = \dot{\Lambda}_m + \dot{\Lambda}_q + H^{\mathrm{T}} W H \tag{35}$$

$$K = \Lambda^{-1} H^{\mathrm{T}} W \tag{36}$$

The data update term $(\dot{P}_d)$ is missing from the covariance equation and the mapping $(\dot{\Lambda}_m)$ and process noise $(\dot{\Lambda}_q)$ terms are missing from the information filter equation and these may be obtained by transformation using matrix identities. For the covariance and information equations, we need the following matrix identities.

$$P\Lambda = I$$

$$\dot{P}\Lambda + P\dot{\Lambda} = 0$$

$$\dot{P} = -P\dot{\Lambda}\Lambda^{-1} = -P\dot{\Lambda}P \tag{37}$$

$$\dot{\Lambda} = -P^{-1}\dot{P}\Lambda = -\Lambda\dot{P}\Lambda \tag{38}$$

Applying these identities to the above matrix differential equations, we have

$$\dot{P} = FP + PF^{\mathrm{T}} + GQG^{\mathrm{T}} - PH^{\mathrm{T}} W H P \tag{39}$$

$$K = P H^{\mathrm{T}} W \tag{40}$$

The covariance filter in this form is called the continuous form of the Kalman-Bucy filter. For the information filter, we have

$$\dot{\Lambda} = -\Lambda F - F^{\mathrm{T}}\Lambda - \Lambda GQG^{\mathrm{T}}\Lambda + H^{\mathrm{T}} W H \tag{41}$$

$$K = \Lambda^{-1} H^{\mathrm{T}} W \tag{42}$$